

Cloudbursting: How do I use EBS volume snapshots to boot my cloud nodes?

Why should I use snapshots?

Because it reduces the provisioning times of your cloud nodes.

Some more background details

By default, all newly-created cloud nodes, and their associated EBS volumes, are empty. This means that the first time a node boots, these volumes have to be provisioned from scratch. That is the case for all cloud nodes (cloud directors, and cloud compute nodes). The provisioning time is most noticeable for cloud directors, as these have to be provisioned over the internet.

Using snapshots for cloud nodes can drastically reduce cloud node provisioning times. This is achieved by pre-provisioning the state of the EBS device with data (the snapshot) that has already been pre-provisioned to AWS.

Using snapshots is highly recommended if you find yourself often terminating your cloud nodes, and then reprovisioning them from scratch.

When done properly (if the snapshot is up-to-date), it can cut down the cloud-director provisioning times, from hours, to minutes. Same applies to regular cloud compute nodes.

This article describes how to:

- create a volume snapshot using AWS web console
- create a volume snapshot using AWS CLI
- use that snapshot to speed up the cloud node provisioning process

Introduction and requirements

This article is for Bright Cluster Manager 7.2. It describes a feature which was released as an update to 7.2, and is only available for CMDaemon revision 31952 and higher. So, to use snapshotting, you need to have a CMDaemon binary of this particular revision, or higher.

Your CMDaemon revision can be checked with:

```
$ cmd -v
```

Cloudbursting: How do I use EBS volume snapshots to boot my cloud nodes?

Sun May 29 13:36:11 2016 [CMD] Info: CMDaemon version 1.8 (r31952)

For snapshots to work, and speed up the cloud node provisioning process, the snapshot which is used needs to be taken from a reference volume that has the same partition layout as the partition layout defined for the cloud nodes (or, specifically, for volumes) that will use that snapshot. Otherwise, the snapshot will still be used to create a new EBS volume, but the node-installer will recreate the partition layout of the EBS volume to match the partition layout that is defined for the node.

By default, all cloud nodes are provisioned using a so-called "sync" install. This means that only the difference between the cloud node's reference software image, and the filesystem already on the cloud node are synchronized to the cloud node. The more similar the snapshot is to the state of the node's software image, the more speedup using a snapshot will provide. This means that, if changes are introduced to the cloud node's software images, it is sensible to create a new, updated snapshot, to include those changes.

For advanced users: note that one can use "exclude lists" to exclude certain parts of software image from being provisioned to the cloud nodes. This might be useful to remember if, for whatever reason, it is desirable to have a reference snapshot that differs significantly from the cloud node's software image. More information on using exclude lists can be found in the admin manual.

Snapshots can be used in both cluster-on-demand and cluster-extension scenarios. This article focuses on cluster-extension, but similar steps can be applied to create and use snapshots for cloud nodes in a cluster-on-demand scenario.

Creating a reference node

Before you create a new snapshot for later reuse, you will need to pre-provision a cloud node. The block device of the cloud node is used as the source of the snapshot.

- provision cloud director without a snapshot (might take a few hours, depending on your uplink speed, local filesystem size, etc)
- stop the cloud director (either via AWS web console, or via cmsh)
- wait 3-5 min for linux to shutdown

At this point you have your reference node and you can proceed with creating the snapshot using either the AWS console, or the AWS cli. Both steps are described next.

Cloudbursting: How do I use EBS volume snapshots to boot my cloud nodes?

Creating the snapshot using AWS web console

- use AWS web console to create a snapshot of director's "/dev/sdb"
 - find the EC2 instance of the director in "Instances" section, select it
 - in the 'description' tab at the bottom, scroll down to "Block device"
 - left click /dev/sdb, context menu will pop up, click on the EBS ID, this will navigate you to "Volumes" menu, with that volume selected
 - right click on the volume, select "create snapshot", fill out the name, click Create
- wait 30-50 min for snapshot to get created (depending on the size of the Director's EBS storage)
- periodically F5 refresh the "Snapshots" list. It will show the progress in 10-15% increments (i.e. it might take a while to suddenly go from 0 to 15%, be patient)
- set the snapshot ID in cmsh (as above) for director, and optionally for chosen cloud nodes
- if the cloud nodes use the same software image as the director, then they can also use the director's snapshot. If the software image of cnodes significantly differs from the director's, then it makes sense to repeat the entire process with a regular cloud node, and to create a separate snapshot for those.

Creating the snapshot using AWS command line tool

In some cases access to AWS web console might not be available. This will be the case if you are using an IAM account with no console access. In such a case, the snapshot will have to be created via the command line. Steps below are a snippets from CentOS 7u1

```
$ yum install python-pip
```

```
$ pip install awscli --ignore-installed six
```

```
$ aws configure
```

```
AWS Access Key ID [None]: <your access key>
```

```
AWS Secret Access Key [None]: <your secret access key>
```

```
Default region name [None]: <your region of preference>
```

Page 3 / 7

(c) 2020 Bright Computing <kb@brightcomputing.com> | 2020-10-24 19:29

URL: <http://oldkb.brightcomputing.com/faq/index.php?action=artikel&cat=12&id=312&artlang=en>

Cloudbursting: How do I use EBS volume snapshots to boot my cloud nodes?

Default output format [None]: <ENTER>

```
$ aws ec2 describe-instances | less
```

This will produce a lot of JSON output.

You need to locate this section for the cloud node which you want to snapshot

```
{  
  
  "DeviceName": "/dev/sdb",  
  
  "Ebs": {  
    "Status": "attached",  
    "DeleteOnTermination": true,  
    "VolumeId": "vol-4d5a63f1",  
    "AttachTime": "2016-05-12T20:29:04.000Z"  
  }  
}
```

```
$ aws ec2 create-snapshot --volume-id vol-4d5a63f1 --description "This  
is my root volume snapshot."
```

```
{  
  
  "Description": "",  
  
  "Encrypted": false,  
  
  "VolumeId": "vol-4d5a63f1",  
  
  "State": "pending",  
  
  "VolumeSize": 42,
```

Cloudbursting: How do I use EBS volume snapshots to boot my cloud nodes?

```
"Progress": "",  
"StartTime": "2016-05-29T10:39:34.000Z",  
"SnapshotId": "snap-3ff26bc7",  
"OwnerId": "137677339600"  
}
```

Monitor progress with:

```
$ aws ec2 describe-snapshots --snapshot-id snap-3ff26bc7
```

```
{  
  "Snapshots": [  
    {  
      "Description": "",  
      "Encrypted": false,  
      "VolumeId": "vol-4d5a63f1",  
      "State": "pending",  
      "VolumeSize": 42,  
      "Progress": "0%",  
      "StartTime": "2016-05-29T10:39:34.000Z",  
      "SnapshotId": "snap-3ff26bc7",  
      "OwnerId": "137677339600"  
    }  
  ]  
}
```

Cloudbursting: How do I use EBS volume snapshots to boot my cloud nodes?

Setting the Snapshot ID using CMSH

Once the snapshot has been created using one of the above-mentioned methods, and once the creation "Progress" is at 100%, you can start using it.

To do so, you need to set the snapshot ID for a EBS object for the nodes which are supposed to use the snapshot. In the example below, we will set it for the cloud director.

If the filesystems of the EBS volume of the cloud director and cloud compute nodes are similar, then the same snapshot can be used for both. The filesystems are considered similar if they use the same (or very similar) software image, and have the same partition layout.

If this is not the case, then if you also want to use the snapshots for cloud compute nodes, you should repeat the process with a cloud compute node.

Here are the steps showing how to set the snapshot ID. This should be done **before** the node is first powered on.

```
$ cmssh
```

```
[headnode]% device use eu-west-1-director
[headnode->device[eu-west-1-director]]% cloudsettings
[headnode->device[eu-west-1-director]->cloudsettings]% storage
[headnode->device[eu-west-1-director]->cloudsettings->storage]% list
Type Name (key) Drive Size Volume ID
-----
--
ebs ebs sdb 42GB
ephemeral ephemeral0 sdc 0B ephemeral0
[headnode->device[eu-west-1-director]->cloudsettings->storage]% use ebs
[headnode->device[eu-west-1-director]->cloudsettings->storage[ebs]]% set snapshotid snap-1a24c8f2
```

Cloudbursting: How do I use EBS volume snapshots to boot my cloud nodes?

```
[headnode->device[eu-west-1-director*]->cloudsettings->storage[ebs*]]%  
commit
```

If you cannot access the "snapshotid" field as shown above, that means you are running an old CMDaemon and/or old CMSH. See the top of the article on how to check the exact revision of CMDaemon that you are running.

Now all you need to do is power on your cloud nodes. By default, if the snapshot's filesystem matches the one described in the disk-setup, the cloud nodes will now use the snapshot to boot.

How do I verify a snapshot is being used?

You can confirm that the snapshot is used properly by SSHing to the cloud node right after it has started its provisioning process, and confirming that the filesystem that is mounted under "/localdisk" contains the data from the snapshot.

Unique solution ID: #1312

Author: Piotr Wachowicz

Last update: 2016-05-30 11:36